



Anderson, I. G. (2018) Understanding the digital legacy of the World War One: Cymru1914. *Cultural Trends*, 27(2), pp. 99-118.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/159061/>

Deposited on: 31 May 2018

Enlighten – Research publications by members of the University of Glasgow_
<http://eprints.gla.ac.uk>

Understanding the Digital Legacy of the World War One: Cymru1914

Ian G. Anderson

Information Studies, University of Glasgow, Glasgow, UK

CONTACT: Ian.G.Anderson Ian.G.Anderson@glasgow.ac.uk

This research paper evaluates the use and impact of the Joint Information Systems Committee (JISC) funded Cymru1914 online digital archive of primary sources relating to World War One (WWI) from Libraries, Special Collections and Archives across Wales. Commemoration of WWI has seen a boom in project funding, including digital activity, funded through disparate initiatives. However, there has been little formal research or analysis of the use and impact of these digital outputs or their implications for funding policy. This research, therefore, seeks to answer two key questions: what is the use and impact of the Cymru1914 digital resources and secondly, what are the implications for evaluation methods and funding policy for similar projects?

Keywords: digital, heritage, methods, evaluation, WWI, Wales

Introduction

In recent decades, the creation of digital cultural heritage collections has become commonplace. Whether from archives, libraries or museums, online digital collections have become established features of our cultural landscape and the sums involved in their creation substantial. It has been estimated that between the 1990s and 2005, £130 million had been spent on digital content creation in the UK (Hughes, 2012, p. 63). Between 1997 and 2011, UK libraries have spent over £100 million on digitisation (Tanner & Deegan, 2011, pp. 8-9). Over the period 2003-2013, The Joint Information Systems Committee (JISC) invested £36.4 million in content digitisation and related projects, including Cymru1914.ⁱ

However, this trend disguises changes in funding. There has been a move away from funding by research councils and the JISC; the Heritage Lottery Fund (HLF) had a long pause in its funding of digitisation projects, and many cultural institutions have engaged in commercial partnerships as well as continuing to fund digitisation from their own resources. Nevertheless, funding for digitisation projects has continued, most recently related to projects to commemorate World War One (WWI), which have seen something of a funding boom. By September 2014, the HLF had awarded £58 million in grants to projects across the UK to explore aspects of WWI heritage, many of these with a digital component.ⁱⁱ To support these HLF-funded projects the Arts and Humanities Research Council (AHRC) awarded £2 million in 2013-2014 for WWI Engagement Centres.ⁱⁱⁱ Each of the five Engagement Centres provides different areas of expertise to support communities undertaking commemoration activities, with the Living Legacies 1914-18 Centre, at Queen's University Belfast, including expertise in digitisation.

Yet, in spite, or perhaps because, of the seeming ubiquity of online digital collections, we know relatively little about their use and impact. There appears to be continual faith in the notion that if you "build it, they will come", making digitisation essentially a supply, rather than a demand-driven model. In 2006, the LAIRAH project (Log analysis of Internet Resources in the Arts and Humanities) estimated that between 30% and 35% of publicly funded digital resources in the arts and humanities went

unused (Warwick, 2006, p 5). The issue is not that we not know how to create successful digital projects - guides by Ross et al. (2002), Lopatin (2006) and Terras (2008) exist alongside many others - but, published evidence of their success, or otherwise, is relatively scarce.

This scarcity of summative evaluation is not due to a lack of appropriate methodology. Since the early Lottery New Opportunity Fund digitisation programme (Brophy & Woodhouse, 2002; Woodhouse, 2001), evaluation and impact methods have evolved (Education for Change, 2006) from their first faltering applications (Anderson, 2007). There is a well-established literature on evaluating the cultural value of heritage, especially in relation to HLF frameworks. These, however, tend to focus on the HLF's generic programmes (Clark & Maeer, 2008) and particular types of heritage (Shipley, et al., 2004), rather than digital projects *per se*. The HLF is not unaware of this. They have a comprehensive guide to evaluation good practice for all types of project ^{iv}(although the recent analysis of 200, 2008- 2013 Heritage Grants projects' self-evaluations reveals major weaknesses in the submitted reports. The authors found that evaluation had not been a prioritised and many lacked evidence or evaluation data (RF Associates, 2017). While HLF issued a tender for a "Review of implementation of digital policy and Terms of Grant for digital outputs" in 2015, no report appears to have been produced to date.

Hughes' work (2012) explicitly sought to fill the gap in evaluating and measuring digital collections and Dobрева et al., note that

In fact, we are currently witnessing a paradox: major institutions from the cultural heritage sector clearly emphasize the place of user evaluation and feedback in digitization-related policies. But in reality, decisions about aspects of digitization that impact users are frequently taken without direct user involvement. (p. 73)

Indeed, it can often be observed that what is presented as "user evaluation" is not in fact evaluation of users; rather, it is using users as a vehicle to validate organisational, operational or technological perspectives.

Sustained development of evaluation methods can be found in parallel initiatives by the JISC and in the field of Digital Humanities. Initial evaluations by JISC of their Phase One digitisation programme (2004-2007) evaluated the impact in terms of the formative contribution of the projects to areas such as standards, metadata and Intellectual Property Rights (IPR) frameworks (Evidence Base, 2007). However, by 2008, JISC had commissioned the Oxford Internet Institute to develop a Toolkit for the Impact Evaluation of Digital Scholarly Resources (TIDSR) since updated in 2011 and 2013.^v

It is perhaps not surprising, given JISC's focus on the academic community both in its content development and support of TIDSR, that the bulk of impact evaluation should focus on scholarly use. Indeed, the 20 or so case studies presented on the TIDSR web site are a testimony to the scale and depth of impact that these digital collections are having on research and scholarship.^{vi} Examining the 29 publications citing the TIDSR toolkit reveals a similar pattern - a preoccupation with metrics, altmetrics, infrastructure and the academic research community.^{vii} This reflects the fact that understanding how digital resources are changing scholarship is perhaps the major research theme of Digital Humanities over the last decade.

More recently, additional methodological approaches have been developed. Tanner's (2012) Balanced Value Impact Model (BVIM) provides a comprehensive analytical framework for impact assessment of cultural resources, particularly the digital. Recently updated (Tanner, 2017), this proposes five stages: Context, Analysis and Design, Implementation, Outcomes and Results, Review and Respond. Each of these stages contains multiple activities, defined variously as steps, components, elements and perspectives. Project-appropriate evaluation criteria are then used to populate a framework to guide and manage the evaluation process. The BVIM has been used by the Wellcome Library to evaluate the impact of their Codebreakers: Makers of Modern Genetics project (Tanner, 2016).

Other approaches have sought to focus on assessing the economic impact of digital content and services, most notably the British Library (Tessler, 2013), although ascribing economic value to such services remains problematic. Meanwhile, the Manchester Metrics Pilot project (Knell, 2014), now called Culture Counts, has sought to assess the quality of outcomes of cultural organisations through triangulating responses from multiple stakeholders. This shifts away from quantitative measurement of outputs and towards a quality metrics framework. This approach is not without its critics (Gilmore, Glow & Johanson, 2017; Selwood, 2016) but has gained attention as a different approach to measuring quality. In higher education, the introduction of impact assessments to the 2014 Research Excellence Framework has concentrated minds on how to measure the impact of research outside of education and prompted new analysis (HEFCE, 2015).

In spite of these methodological innovations, there still remains a gap between the need to better evaluate the use and impact of more public facing digital heritage projects funded by the likes of the HLF, and the application of methodological expertise built up in the Digital Humanities, Information Studies and Cultural Heritage sector. It is this gap that this research attempts, in part, to fill and in so doing highlight methodological issues and policy implications.

Methodology

The decision to select the Cymru1914 site for evaluation was taken based on several factors. Although JISC funded, Cymru1914 was included within the remit of Living Legacies 1914-18 Centre as it has by far the largest volume of digital outputs, comprising 1,292,013 individual items arranged into seven categories, as shown in Table 1. Material was sourced from archives, libraries and special collections of local authorities, universities and records offices across Wales.

Table 1. Type and Volume of Digital Resources

Type	Volume
Newspapers	1,129,096
Archives	142,292
Manuscripts	13,641
Journals	5,424
Photographs	1,076
Photo Albums	450
Sound	34

Cymru1914 states its purpose thus:

The project will make available a coherent, consolidated digital collection revealing the often hidden history of the First World War as it impacted all aspects of Welsh life, language and culture. This digital archive brings together source materials that were previously fragmented and frequently inaccessible. This digital archive is a unique resource of vital interest to researchers, students, and the public in Wales and beyond. (cymru1914.org)

The methodology adopted for this research is based on the TIDSR toolkit.^{viii} In this case, web analytics, in the form of Google analytics, online user survey and user interviews were used. TIDSR is not prescriptive in relation to the instruments that should be used; rather, it provides a suite of descriptions of relevant methods and case studies divided between quantitative and qualitative.

There are a number of reasons why TIDSR was selected. Firstly, the original grant application to AHRC as part of the Living Legacies 1914-18 project specified TIDSR as the methodology to be adopted in the research. Subsequently, consideration was given to other approaches. BVIM does provide a fuller critical framework for impact assessment and as TIDSR was used to populate its methods appendix, compliance with methods specified in the Living Legacies grant application could have been maintained. However, the comprehensive nature of BVIM brings with it considerable complexity - its documentation running to some 119 pages. Whilst some aspects of BVIM are relatively easy to complete, such as context, given the time and resources available it was deemed unrealistic to adopt the model. It was considered preferable to adopt the simpler TIDSR approach, which also had the benefit of a larger body of completed case studies. Whether this was the correct judgement to have made will be returned to later. Assessing the economic impact of Cymru1914 was outside the scope of the research, and it was deemed unviable to adopt methods that required face to face meetings, workshops or focus groups, such as Culture Counts.

Access to the Cymru1914 Google analytics was provided by the National Library of Wales (NLW) and data were downloaded for the period from 1 April 2016 to 31 March 2017. This provided a sufficiently long period to get a sense of traffic and usage trends. This period also encompassed the time that the online user survey was initially available so data from the two would have contemporaneity. In total, 13 types of analytic data were collected. One drawback to the use of this particular data set is that data will not be collected if a user blocks cookies on their computer. This problem is not unique to Google analytics or other web analytic tools. Whilst providing a good view of overall site usage, web analytics are less useful for gauging impact.

To this end, a user survey was developed to provide a fuller picture of who the users were, how they were using the site and the impact it was having. As with any survey, a balance needed to be struck between gathering sufficient, relevant information and length of survey. Based on prior experience of online surveys, it was judged that 10 to 15 questions were optimum and would not induce survey fatigue. The full survey, which comprises 13 questions, can be found in the appendix. The user groups identified in question 4, purposes in question 5 and subject areas in question 6 were based on those envisaged in the original grant application by the NLW to the JISC for the Cymru1914 project. Questions 7 and 8 on user actions and recommendations are attempts to gauge impact. The survey was reviewed by NLW staff and minor changes to wording made before being validated. The survey was hosted by NLW on the online survey site, SurveyMonkey, and linked to the home page of the Cymru1914 site. The survey was available in English and Welsh. It went live on the 10th November 2016 and ran for 6 months until 11th May 2017. Although it was originally intended to run until the middle of March 2017, the survey period was extended for two months to increase participation. This had limited effect, however, with only five further responses being received during this period despite a renewed round of promotion.

The survey had promotion targeted at identifying users of the Cymru1914 site. It was highlighted to NLW friends' groups by the NLW and via Twitter to academics with Welsh history interests. In all cases potential respondents were directed to the Cymru1914 site, rather than the SurveyMonkey site. The reason for this approach was to encourage responses from genuine users of Cymru1914, rather than to simply generate survey responses per se.

In total there were 47 responses, a very low participation rate. There were no incomplete surveys, which suggests that the survey itself was not too onerous. Given that the site attracted approximately 21,000 users during the period the survey was online, the response rate was 0.2%. With a 95% confidence level, this gives a margin of error of 14%. To get a more typical 5% margin of error (for the same 95% confidence level) 378 survey responses would be needed. This does not seem an unrealistic target but given the seeming reluctance of Cymru1914 users to complete the survey even doubling the number of responses to 100 would still leave a 10% margin of error. Low response rates to online surveys are typical and comparable projects, such as the

evaluation of the SPHERE, an even lower response rate of 11 was returned^{ix}. A low response rate does not mean the results are not useful and caution always needs to be exercised with self-completed surveys. As can be seen in the impact section below, the survey does capture most of the user and usage types envisioned by the Cymru1914 project. Bearing in mind the margin of error, the survey results should at least reveal some broad trends.

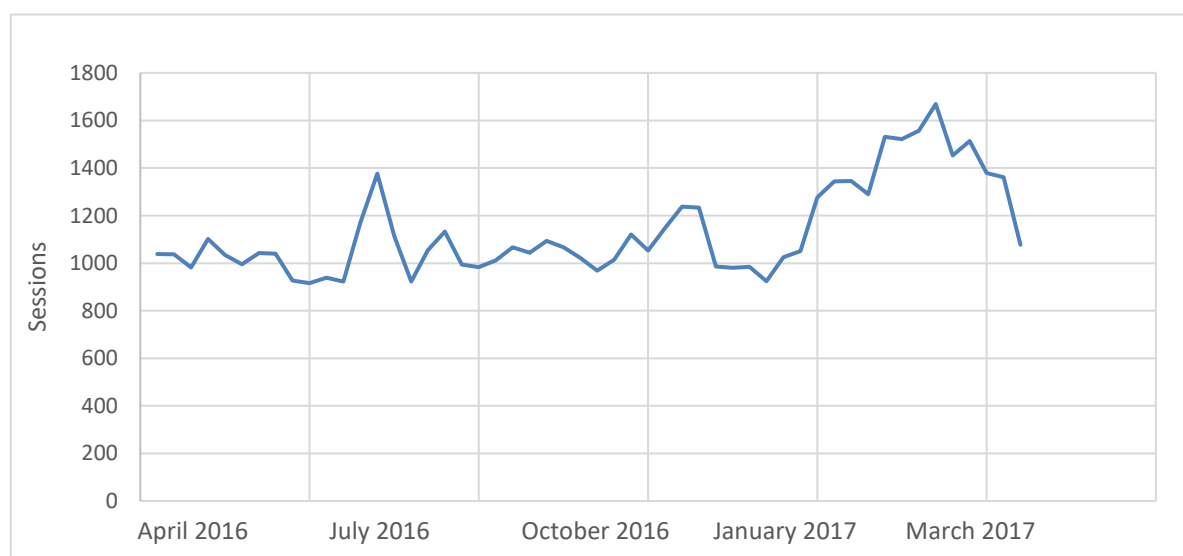
One problem that was only noticed once the survey data were received was that the way in which response options had been formatted allowed respondents to select multiple responses to some questions rather than single ones. This was because check boxes rather than radio buttons had been used. Respondents could select more than one user group, purpose and subject area (questions 4, 5 and 6). In looking at the results it became clear that some respondents were genuinely expressing multiple roles. For example, one respondent selected academic and general public user groups, but also academic and personal research purposes. As these results appeared to be expressing a genuine pattern of user behaviour, rather than the mistaken clicking of boxes, it was decided to leave these dichotomy responses intact.

The web survey asked respondents to provide an email address if they were willing to participate in a follow up interview. Five respondents provided email addresses and two of these agreed to be interviewed. Interviews were semi-structured and asked participants to describe their research, how they came to use the Cymru1914 site, what they were looking for, how they used what they found, the value they thought the site provided and if they were involved in any commemoration activity. Interviews were conducted over the telephone and were not recorded, but notes were taken during the conversation.

Use

Overall usage of the Cymru1914 over the 12-month period based on Google analytics is 43,592 users, covering 59,390 sessions (Google analytics uses the term sessions rather than visits) and 310,440 page views. The trend in sessions can be seen in Figure 1.

Figure 1. Session Trend April 2016 – March 2017



A noticeable peak in July 1916 coincides with the anniversary of the beginning of the Battle of the Somme. There is also a slight upturn in November when the Somme offensive ended. Another, more prolonged rise can be seen from January 2017 that seems to follow the publication of Jonathan Scott's *Tracing Your British and Irish*

Ancestors in October 2016, which references Cymru1914. There is certainly no WWI event or other publicity to explain the increase. The visitor numbers for the site certainly appears healthy, but without the baseline of a similar project it is hard to say how successful it has been. To put this in context, the NLW web site (the parent of Cymru1914) attracts between 45,000 and 170,000 visits per month, compared to an average of 4949 per month for Cymru1914. Nevertheless, Cymru1914 is certainly able to attract new visitors, with 72% of all sessions over the period coming from new users. What remains to be seen is how sustainable these numbers are beyond the commemoration period. Clearly there will always be a residual interest in WWI but what we do not know is how much of the current traffic is due to short-term interest. The risk for any "event" type digitisation projects is that it will have a limited lifespan.

Beneath these headline figures is the noticeably high bounce rate of 61% - bounce rate being the percentage of single page visits.^x This reveals the proportion of visitors who left a site from the page they entered without any other interaction. A high bounce rate is not necessarily problematic. This may well indicate that users arrive on exactly the page they were looking for and leave, particularly for content sites such as Cymru1914, which can often have 40-60% bounce rates.^{xi} However, it can also indicate that users web search terms, which whilst directing to the site, does not provide the content they were looking for. As web analytics expert, Avinash Kaushik states, "anything over 35% is a cause for concern and anything above 50% is worrying".^{xii} As 79% of Cymru1914's traffic comes from organic search (search engine searches, not paid search) compared to 15% from direct links, this is worth further investigation.

Not surprisingly, Google is the biggest source of visits to Cymru1914. Seventy-six percent of all sessions which originate with it have a bounce rate of 63%. Searches through Google on the android mobile platform and Facebook also post high bounce rates at 60%. However, the analytics also show these users viewing an average of 4.5 pages per session, and averaging a session duration of 3 minutes 13 seconds, as can be seen in Table 2.

Table 2. Sessions by Acquisition Channels

Source / Medium	Acquisition			Behaviour		
	Sessions	% New Sessions	New Users	Bounce Rate	Pages / Session	Avg Session Duration
	59,390	72.09	42,812	61.20	5.23	00.03.31
google / organic	45,407	72.87	33086	63.35	4.50	00.03.13
(direct) / (none)	8,837	69.42	6135	55.51	6.39	00.03.59
bing / organic	1,245	67.87	845	45.14	14.93	00.07.02
com.google.android.googlequicksearchbox / referral	793	84.99	674	67.97	3.11	00.01.44
yahoo / organic	377	77.72	293	51.99	8.36	00.07.32
m.facebook.com / referral	340	92.94	316	66.47	2.22	00.01.04
discover.library.wales / referral	183	50.82	93	39.34	5.74	00.04.43
t.co / referral	159	57.86	92	53.46	7.87	00.02.49
llgc.org.uk / referral	132	34.85	46	31.06	9.09	00.03.31
walesatwar.org / referral	126	39.68	50	50.00	9.77	00.03.31

This suggests that underneath these high bounce rates there is still an appreciable level of engagement beyond a single page. Although the proportion of users coming from the Bing and Yahoo search engines is far smaller, their lower bounce rates and greater number of page views may suggest that there is something in the way Google handles searches or the type of user who prefers Bing and Yahoo over Google. However, a more detailed analysis than can be gleaned from the aggregate analytics available would be required to explore those hypotheses.

As Table 2 indicates referral sites, those that provide direct links to Cymru1914, provide the lowest bounce rate of between 30% and 50% (with the exception of Facebook), and the higher range of page views per session, although there is more variation on the session duration. This ranges from 2.49 minutes from Twitter to over 10 minutes from the Wales at War website.

Looking in more detail at Google analytics' engagement measures shows more clearly the relationship between session duration and pageviews in Table 3.

Table 3. Pageviews and Session Duration

Session Duration	Sessions	Pageviews
0-10 seconds	37,214	38,201
11-30 seconds	2,272	5,692
31-60 seconds	2,706	8,587
61-180 seconds	5,544	26,891
181-600 seconds	5,931	54,654
601-1800 seconds	4,156	74,437
1801+ seconds	1,567	101,978

Although the majority of pageviews last less than 10 seconds, this is not necessarily problematic, provided one assumes that this is part of the process of users skipping in and out of information to find what they are looking for. Given the large number of newspaper sources this could be possible. What is also notable is the increase in the number of sessions lasting longer than one minute compared to sessions lasting 31 seconds to one minute and a corresponding increase in page views. Of note are the number of sessions lasting longer than 10 minutes and the high number of pageviews. This suggests that there is a small, but intensive, group of visitors making use of the site.

Looking at measurements of page depth (the number of pages viewed in a single session) adds to the notion of intensive use. As Table 4 illustrates, there is a proportional decline between page depth, sessions, and pageviews from a single page down to 19 pages, then a large spike for page depth above 20 pages. Although the number of sessions viewing 20 or more pages is small at 5.4% of total sessions, these account for 55% of total page views.

Table 4. Page Depth, Sessions and Pageviews

Page Depth	Sessions	Pageviews
1	36,347	36,347
2	7,095	14,190
3	3,445	10,335
4	1,799	7,196
5	1,408	7,040
6	941	5,646
7	826	5,782
8	615	4,920

9	612	5,508
10	454	4,540
11	440	4,840
12	367	4,404
13	377	4,901
14	254	3,556
15	317	4,755
16	236	3,776
17	253	4,301
18	199	3,582
19	190	3,610
20+	3,215	171,211

The notion of a small group of intensive users is supported by looking at the frequency of sessions in Table 5. Once again, there is a sharp decline in the number from single sessions down to eight sessions, then a sharp increase in the 9-14 session band, before steadily declining again.

Table 5. Frequency of Sessions and Pageviews

Count of Sessions	Sessions	Pageviews
1	42812	161627
2	5808	35319
3	2268	16036
4	1248	10498
5	818	6767
6	602	6610
7	478	5456
8	382	4301
9-14	1437	19045
15-25	1179	16385
26-50	1046	14888
51-100	671	7528
101-200	437	3843
201+	204	2137

Perhaps not surprisingly, almost 83% of users were from the UK, followed by the US at 4%, Australia 2%, Canada 1.5%, India 1.25% then France, Germany, New Zealand and Ireland all at less than 1% of total visitors. Frustratingly, Google analytics cannot determine the proportion of users from Wales. Interestingly, the bounce rate for the UK drops to 58%, with the rate for all other countries except New Zealand above 63%. Cymru1914 certainly seems to have the global reach it hoped for, the site has been accessed from almost every country in the world, even if overseas use is very limited.

Turning to the use of the search feature in Cymru1914, 14.4% of visits, or 8528 sessions, used the search function. As the total of unique searches is over 19,000 (twice the number of sessions using search), this indicates a number of sessions using "search" more than once with different terms. There is no "right" level of "search" - high search rates can mean users are finding it hard to navigate a web site, whilst low rates can mean users cannot find the search function or find it confusing. Of more importance is

search quality. If the percentage of search exits (searches which are made immediately before leaving the site) are combined with search refinements (where users searched again immediately after a search) and deducted from total searches, we are left with 63% of searches being “successful”. These search results are followed by one or more pageviews, with users on average viewing 3.48 pages. Google analytics provides the top-10 search terms. Wildcard searches registered over 5% of total unique searches, the rest of the top-ten all registered below 1%. Server log files would provide more details on search terms, but were not available for this research.

Survey results can add further details about usage. Ten respondents, or 21%, completed the survey in Welsh. This is somewhat higher than the 15% of the population able to write in Welsh according to the Welsh Language Use Survey 2013–14 (Welsh Government, 2015, p 35).

There was a higher proportion of survey respondents above 35 years of age than the Welsh population, and smaller proportion of respondents younger than 35. No under 18s responded to the survey. The age profile of respondents is shown in Table 6.

Table 6. Age profile of Survey Respondents v Welsh Population

	Survey Percent	Population Percent
65+	23	18
55-64	19	13
45-54	19	14
35-44	15	14
25-34	11	12
18-24	6	11
Under 18	0	18
Total	93	
Missing	7	
Total	100	100

Judging by the survey evidence, Cymru1914 can attract users across the range of its anticipated user groups, as seen in Table 7. Note that this is one of the questions where users could select multiple answers, so the total is greater than 47. The only user groups not picked up by this survey were government and military employees.

Table 7. User Groups

	Frequency	Percent
Academic	11	18.3
Library & Archive Staff	8	13.3
Student (HE)	3	5.0
Schools (Teacher or Pupil)	2	3.3
Government	0	0.0
Museum Staff	2	3.3
General Public	25	41.7
Technical / e-Research	6	10.0

Creative Industries / Business	1	1.7
Media	2	3.3
Military	0	0.0
Total	60	100.0

Respondents were found to be using the site for a wide variety of purposes and subjects that correspond with Cymru1914's objectives, as can be seen in Tables 8 and 9.

Table 8. Purpose Frequencies

		Frequency	Percent
Primary purpose for using Cymru1914	Personal research	27	47.4
	Schoolwork / Homework	0	0.0
	Teaching (primary / secondary)	0	0.0
	Teaching (continuing / further / higher)	2	3.5
	Academic Research	13	22.8
	Journalism / media / creative content	3	5.3
	Community outreach / exhibitions	3	5.3
	eGovernment	0	0.0
	Reference (Library & Archive)	4	7.0
	Other	5	8.8
	Total	57	100.0

Cross tabulating these results is problematic given that some respondents identified with more than one user group, purpose or subject. Nevertheless, experimenting with various forms of cross-tabulation showed some clear patterns. These must be interpreted with caution, but the most pronounced patterns suggest some credibility, if only for being unremarkable. For example, 60% of the general public used the site for personal research and 54% of academics used the site for academic research. Personal research was also the most frequent purpose amongst library and archive staff at 35%.

The noticeable absence, at least in this survey, is the site's use for schoolwork or primary or secondary teaching. Although there were two teachers in the user group, they were using the site for personal and academic research, media and outreach. This suggests the benefit of having separate questions for user groups and purposes as one might (mistakenly) assume that teachers were using the site for teaching purposes. Nor were there any use by eGovernment, at least that the survey shows. The "other" category included two book authors, one "professional (non-academic) research", one "WWI project", and one complaint from a genealogist that they could not get the site to scroll using Chrome on their Android tablet!

Table 9. Subject Frequencies

	Frequency	Percent
--	-----------	---------

Subject Areas	Genealogy / Family History	21	22.1
	Local History	19	20.0
	Welsh History	16	16.8
	Welsh Language / Culture	3	3.2
	Military History / Commemoration	11	11.6
	History of WWI	21	22.1
	Tourism	1	1.1
	Other	3	3.2
	Total	95	100.0

The subject frequencies show a spread across all available options with an unsurprising concentration in genealogy and WWI, closely followed by local history and Welsh history. Although three users checked the "other" option for subject there were actually four responses left. The most revealing was "I use the Cymru1914 site as a teaching aid for a primary-source based module that I teach in a university. It's been a great source of ideas for both student essays and student dissertations." Other responses included barddoniaeth (poetry in English) and the history of agriculture and society.

Looking more closely at how the Cymru1914 has been used, 18 respondents (or 38%) said that they downloaded a resource. However, when one looks at the breakdown of which resources were downloaded, 20 respondents selected newspapers, indicating two more respondents actually downloaded content than had initially said so. Thus, the percentage figure for users who download increases to 42%. The breakdown of downloaded resources is provided in Table 10.

Table 10. Downloaded Resources

		Frequency	Percent
Which resources have you downloaded	Newspapers	20	50.0
	Archives & Manuscripts	9	22.5
	Photographs	7	17.5
	Journals	4	10.0
	Sound	0	0.0
	Total	40	100.0

In one important respect, respondents to the survey are very different from the profile of users revealed by Google analytics. Only 23% of users became aware of the site through a search engine; the largest proportion, 25% by a reference from another source. Finding the site through a search engine was the most common result for the general public, however, with nine respondents (36%) finding the site this way, followed by seven (28%) who found a reference in another source and four (16%) who were notified through a mailing list. The full breakdown is in Table 11.

Table 11. How did you become aware of Cymru1914?

	Frequency	Percent
Mailing list	4	8.5

Personal email message	3	6.4
Reference from another source (e.g. website, forum)	12	25.5
Search engine	11	23.4
Word of mouth	6	12.8
I was involved in the work of the site	6	12.8
Social media	1	2.1
Other (unspecified)	2	4.3
Missing	2	4.3
Total	47	100.0

Users were asked to rate overall how easy they found the site to use, and the clear majority, 74.5%, found the site quite easy or very easy. Users were also asked to rate individual features of the site on a 1-5 Likert scale. Using Cronbach's Alpha to test the reliability of the scale produced a result of 0.915, anything above 0.7 denotes good reliability, (i.e. that all items in the scale measure the same concept, it is internally consistent.) With 1 representing the lowest score and 5 the highest, the itemised means in Table 12 show that all items in the scale rated quite well with users. Image Zoom/Rotate and Site Information were rated slightly higher than other features.

Table 12. User Rating of Site Features

	Mean	Std. Deviation	Frequency
Search Interface	3.6	1.140	35
Search Results	3.6	1.087	35
Image Zoom / Rotate	3.8	0.964	35
Further Information	3.2	1.555	35
Filter by	3.3	1.132	35
Browse Interface	3.5	0.919	35
Navigation	3.5	1.039	35
Page Layout	3.6	1.193	35
Site Information	3.7	1.120	35

Impact

Having a satisfied user base is, however, only one part of a successful digitisation project. The ultimate purpose of any project should be that it makes a positive difference to users. So, this research also attempted to assess the impact that Cymru1914 was having. This was achieved in part through the user survey, questions 7 and 8, but also through follow up interviews.

Aside from the usual caveats about self-completed surveys, there is the added issue that respondents are being asked to reflect on causation, a difficult thing to do and where caution in interpreting the responses is required. Certainly, some respondents were unable or unwilling to answer this question, six left it blank - a higher rate than for other questions. Ultimately, trust needs to be placed in the answers provided. Looking at the responses there was sufficient variation to suggest that respondents were not simply taking part in a tick box exercise.

Table 13 shows the number of responses for each answer; the percentage given refers to the total number of respondents to the survey (n=47) not the percentage of total answers (n=115) or respondents of this question (n=41).

Table 13. Impact Responses

	Frequency	Percent
Find the information you were looking for	30	63.8
Discover new sources	30	63.8
Change your view / opinion on a subject	6	12.7
Become interested in a new subject	10	21.2
Recommend the site to someone else	18	38.2
Visit an online archive, library or museum site	12	25.5
Visit a physical archive, library or museum	5	10.6
Visit a place mentioned on the site	4	8.5
Missing	6	12.7
Total	115	

Analysing the results more closely, 63.8% for “Finding the information you were looking for” and “Discover new sources” can be considered a very satisfactory, if somewhat elementary, impact. Although the figures for both look suspiciously similar, examining individual responses shows that eight who did not find the information they were looking for, nevertheless discovered new sources. In contrast, seven of those who found what they were looking for did not go on to discover new sources. A smaller, but encouraging, 21% of respondents became interested in a new subject, illustrating the potential of online digital resources to expand people’s horizons, even one with as narrow a focus as Cymru1914.

It was, perhaps, not surprising that “Change your view / opinion on a subject” scored quite low. This is a hard thing to recognise and may occur over a period of time. Nevertheless, there is evidence that Cymru1914 had this impact on some users.

The remaining responses cover what can be considered the higher level impacts, those that result in an action. When online, recommending the site to someone else or visiting another online resource is an action with a low threshold and these rank third and fourth, respectively. The “Recommend the site to someone else” option had a follow up question (question 8) asking for more details. The range of responses is in Table 14.

Table 14. Recommendation Responses

	Frequency	Percent
Family member	12	16.2
Friend	19	25.7
Local history / Genealogy member	14	18.9
Community group member	6	8.1

Teacher	7	9.5
Academic	10	13.5
Other	6	8.1
Total	74	100.0

Although it was intended that this question would only be answered by those who selected the recommend option in question 7, it was in fact answered by 35 respondents, compared to only 18 who selected the option in question 7. Whether you would recommend something to someone else is often seen as a telling endorsement, so these responses constitute a bonus. The “other” category included students, colleagues, researchers, “Dilynwyrr ar Twitter” (Followers on Twitter) and one that would not recommend it “because the discovery interface is awful”.

The remaining two impacts, visiting a physical archive, library or museum or a place mentioned in Cymru1914, are the highest level of impact. These two impacts have the lowest frequency with 10.6% and 8.5%, respectively. The overall level of impact looks impressive with 115 impacts from 47 respondents, almost 2.5 per person.

The last impact question related to the use of social media links within the website. Cymru1914 provided social media links on each page to Facebook, Twitter and Google+. Thirty-eight percent, or 18 respondents, used these links and the breakdown is provided in Table 15. This shows little use of Google+, twice the use of Twitter over Facebook, with six respondents using both Twitter and Facebook.

Table 15. Social Media Link Usage

	Frequency	Percent
Twitter	16	64.0
Facebook	8	32.0
Google+	1	4.0
Total	25	100.0

It is evident that survey respondents are not the one-time “bounce” users that make up the majority of visits indicated by Google analytics. Only four respondents had not used the site before; eight users had visited a few times a year; 16 users used the site monthly; 17 users daily or weekly. (There were two non-respondents.) This provides a glimpse of the more intensive users, who broadly aligned with the uses and users envisaged by the Cymru1914 project. As the project did not benchmark the volume or depth of use or detailed impacts, it is difficult to judge the analytics and survey results. Clearly, there is something of a bipolar distribution, with a high volume of low-intensity use (perhaps even one-time use) and a far smaller, but far more intensive user group.

Although there is reasonable survey evidence that Cymru1914 is having a positive impact for some users, the survey cannot provide insight into how the site fits into users' research, its value to them and influence on their understanding of the Welsh experience during WW1. It is this insight the interviews were intended to provide.

Two survey respondents were willing to be interviewed. This is a small sample, even amongst the number of those responding to the survey, let alone the overall user community. It is possible that the interviewees volunteered because they believed that they have something exceptional to share. However, the purpose of the interviews is not to gain a representative sample, but to develop a more detailed picture of survey respondents' use of Cymru1914.

In the first interview (Subject A) it became clear that Cymru1914 fitted into a long-term interest in WW1, particularly as it related to Caerphilly. The subject's interest began over 20 years ago with visits to the Ypres battlefields, enhanced 15 years ago with a remembrance service for schools, at which the story of one of the soldiers was read out. This was followed by school trips to the battlefields where different themes

were developed, such as the stories of the first recruits. These experiences developed into a broader personal interest that includes collecting memorabilia, especially medals, matching stories to names on medals and those on the local cenotaph. These stories have been shared with local groups such as Guides, Scouts and the British Legion, as well as posting information on online sites such as the Great War Forum and South Wales Militaria Society.

Cymru1914 importance resided in the "sheer convenience of the resources" (Subject A, Interview, Nov 2017) in helping to link soldiers' names with their stories. The ease of access, ability to search by name and manageable scale of the resources were all mentioned as valuable features of the site. Particular appeal lay in these "making it personal" and "bringing it alive" especially with the "vivid writing of the period" (Subject A, Interview, Nov 2017). An example was provided of being able to track down a 93-year-old relative of the Chair of Caerphilly Ruby Club, who had been wounded during the War. The descendant was able to identify faces in period photographs and put names to them. Local people have since been in contact asking for help to research their relatives. Subject A emphasised the importance of storytelling but also linking the virtual research with the physical places. He also mentioned how the War transcends eras, how respectful children could be, how they make a connection with relatives and places to the extent some have been physically moved.

In this case, Cymru1914 has played a relatively small and recent role in the subject's interests, but it was regarded as important in broadening and deepening powerful stories connecting current and future generations with their place in the Great War. Ultimately, the subject has an ambition to write a book on Caerphilly in WW1 for which Cymru1914 will continue to provide a source for stories and inspiration.

The second interview (Subject B) tells a very different story of involvement with Cymru1914. Unlike Subject A, Subject B had no longstanding interest in WW1, no historical background or training and no resources.

Originally, the local tourist office in Rhyl had planned a small exhibition to mark the start of the War. This had been prompted by a local history group article on the "North Wales Pals" (a locally raised formation which became the 13th Battalion, Royal Welsh Fusiliers) who had lined the streets of Rhyl when some Belgian refugees arrived. The focus of the story was the soldiers, not the refugees, and there appeared to be no information on them. No records had been kept locally. County boundaries had changed and records at the National Archives in Kew were incomplete. It was by chance that Subject B was searching online in English and Welsh and entered the term "Cymru 1914". On the site, Subject B found the name of one refugee who had died during their stay and was buried in an unmarked grave. Finding that one name on Cymru1914 site set off a remarkable story: "Without Cymru1914 I would have found nothing. One name led to 66 people, told a story nobody knew about and united families in Belgium who had never met" (Subject B, Interview, Nov 2017).

Three years later, there is only one name left to find from all the refugees who arrived in Rhyl 100 years ago. All their stories of their time in Rhyl were shared with members of the local refugee committee. There have been public talks, three visits to Belgium and the descendants of two refugee families who lived in the same Belgium town, but knew nothing of each other's existence, reconnected on a visit to Rhyl. There is now a town plan for commemoration and there have been radio and TV appearances.

For Subject B, the key value in Cymru1914 has been in being able to piece together a good story related to the War. How despite losing relatives during the conflict, members of the committee kept working for the refugees, how there was never any animosity towards them, just as there is none towards refugees in the town today. In Subject B's words, "Rhyl was built as a holiday town to welcome people, it is still welcoming people today" (Subject B, Interview, Nov 2017).

Of the many thousands of users, the experience of Subjects A and B are likely to be outliers, but may not be entirely unique. Their use of Cymru1914, as displayed in the

survey data, is not atypical, although the survey has not captured all instances of such use.

Looking at the overall picture raises the issue, for both funders and projects, about what types of use are considered preferable. High volume, but with less impact, or small volume with higher impact? Is investment best served by making a big difference to a small number of people or meeting the basic information needs of a large number? These are not, of course mutually exclusive, but the resources available for any project are finite. Were the NLW to develop Cymru1914 further, they could add more content or develop more sophisticated search facilities and personalisation features. Neither of these choices would be “wrong” but they would serve different types of users and uses.

Conclusion

This research described here aimed to provide a comprehensive evaluation of a non-academic digital resource, covering use and impact using established methods and tools. Although not perfect, the results have provided useful insights into this approach and provided a benchmark for evaluating similar projects. It has shown that Cymru1914 can have a deep impact on users, although this represents the one area where more individual stories would be beneficial. Such stories provide the most compelling picture of the physical and virtual connections that link people, places and their heritage.

Had the research adopted the BVIM it is likely that extra dimensions would have been added to the impact assessment. However, even if the research had the resources to say, contact all the content providers across Wales, as well as institutional stakeholders, etc. BVIM would not have addressed the main problem of low survey responses, although it may have mitigated its effect. TIDSR seems well suited to academic related resources where there is clearly defined user group or groups, where demographic profiles are easier to establish for benchmarking and whose interests are closely aligned with the resource. It would also seem suitable where data on particular key performance indicators are required.

One recommendation for funders would be to require applicants to provide more detail on expected impacts, especially regarding usage levels, type of engagement and success criteria. Looking at the impact requirements in UK research council applications may be useful in this respect. It would certainly make evaluating the impact easier. Whether it would be enough to nudge HLF grantees who are unwilling or unable to complete better evaluation reports remains doubtful, especially given that these reports are already required by the HLF and there is a range of guidance available. In this respect funders may need to consider conducting impact evaluations themselves, or commissioning third parties to do so.

Adapting something like the Culture Counts approach to digital resources, and virtual audiences, would be challenging, but could be a useful complement to existing approaches irrespective of the assessment framework, methods adopted or who conducts the evaluation. Low survey responses remain a problem that that can undermine even the most rigorous approach. Other than offering inducements to complete surveys, there is little else this research project could have done to increase responses without jeopardising the credibility of the returns. The experience of this research might not be replicated elsewhere, of course, but survey fatigue, particularly for resources aimed at the “general public” is a recognised phenomenon. This may point to less reliance on survey instruments, or at least more targeted deployment, but this would require a better benchmarking in the first place. Certainly, more research in this area would be welcome and could yield the most improvement in our understanding of the impact of digital resources.

Ian Anderson is Senior Lecturer in Information Studies in the School of Humanities, University of Glasgow. His research interest is in digital heritage, in particular evaluating

the creation and use of digital content and online information systems in archives, libraries and museums.

Acknowledgements: The author would like to acknowledge Prof. Lorna Hughes for initiating this project and her support and advice, Dr Leonidas Konstantelos for additional data gathering and analysis, Sian Lloyd Pugh and Dafydd Tudur of the National Library of Wales for their cooperation and lastly the users of the Cymru1914 web site for sharing their experience.

References

- Anderson, I.G. (2007). Pure dead brilliant? Evaluating The Glasgow Story digitisation project. *Program: Electronic Library and Information Systems*, 41 (4), pp. 365-385.
- Brophy, P. and Woodhouse, S. (2002). Evaluation and Impact assessment for NOF Digitise Projects. An Information Paper from the NOF Technical Advisory Service., NOF-digitise Technical Advisory Service.
<http://www.ukoln.ac.uk/nof/support/help/papers/impact-assessment/>
- Clark, K & Maer, G. (2008). The cultural value of heritage: evidence from the Heritage Lottery Fund, *Cultural Trends*, 17 (1), 23-56.
- Education for Change (2006). The Fund's ICT Content Programmes Final Evaluation Report, Big Lottery Fund,
http://www.biglotteryfund.org.uk/er_eval_ict_final_rep.pdf
- Evidence Base, (2007). *Evaluation of the JISC Digitisation Programme Phase 1 and International Contextualisation*. JISC, Bristol.
- Gilmore, A., Glow, H & Johanson, K. (2017). Accounting for quality: arts evaluation, public value and the case of "Culture counts". *Cultural Trends* 26, pp. 282-294.
- HEFCE, (2015). *The nature, scale and beneficiaries of research impact: An initial analysis of Research Excellence Framework (REF) 2014 impact case studies*. Research Report 2015/01, King's College London and Digital Science.
- Hughes, L. M. (2012). *Evaluating and Measuring the Value, Use and Impact of Digital Collections*. Facet Publishing, London.
- Knell, J. (2014.) *Manchester Metrics Pilot Final Report of Stage One*, Arts Council England, Manchester. <http://www.artscouncil.org.uk/publication/manchester-metrics-pilot>
- Lopatin, L. (2006.) Library digitization projects, issues and guidelines. *Library Hi Tech*, 24(2). 273-289.
- RF Associates. (2017). *Heritage Grants: A Review of the Self-Evaluations and Outcomes of 200 completed projects funded during Strategic Plan 3. Final Report*. Heritage Lottery Fund, London.
- Ross, S., Anderson, I., Duffy, C., Economou, M., Gow, A., McKinney, P. and Sharp, R. (2002). *The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials*. National Initiative for a Networked Cultural Heritage, Washington DC, USA.
- Scott, J. (2106). *Tracing Your British and Irish Ancestors: A Guide for Family Historians*, Pen & Sword, Barnsley.
- Selwood, S., (21016). Response, *Cultural Trends*, 25, pp. 69-71.
- Shipley, R., Reeve A., Walker, S., Grover P., & Goodey, B. (2004). Townscape Heritage Initiatives Evaluation: Methodology for Assessing the Effectiveness of Heritage Lottery Fund Projects in the United Kingdom. *Environment and Planning C: Government and Policy*, 22 (4), 523-542.
- Tanner, S. G. and Deegan, M. (2011). *Inspiring Research, Inspiring Scholarship: The Value and benefits of digitized resources for learning, teaching, research and enjoyment*. HEFCE, London, pp 8-9.
- Tanner, S. G. (2012). *Measuring the Impact of Digital Resources: The Balanced Value Impact Model*. King's College London.
www.kdcs.kcl.ac.uk/innovation/impact.html

- Tanner, S. G. (2016). Using Impact as a Strategic Tool for Developing the Digital Library, *Library Leadership and Management*, 30 (4).
- Tanner, S. G. (2017). *The Balanced Value Impact Model Version 2 overview graphics in full resolution*, Digital or Visual Products, King's College London.
https://kclpure.kcl.ac.uk/portal/files/80251677/BVI_Model_Version2_graphics_by_Simon_Tanner.zip
- Terras, M. (2008). *Digital Images for the Information Professional*. Ashgate, London.
- Tessler, A. (2013). *Economic valuation of the British Library*. Oxford Economics, Oxford.
- Warwick, C., Terras, M., Huntington, P., Pappa, N., Galina, I. (2006) *LAIRAH project: log analysis of Internet resources in the arts and humanities. Final Report to the Arts and Humanities Research Council*. University College London.
- Welsh Government, National Survey for Wales, 2013-15: Welsh Language Use Survey
- Woodhouse, S. (2001). The People's Network and the learning revolution: building the NOF digitise programme, *Ariadne*, 29.

Notes

- i (<http://webarchive.nationalarchives.gov.uk/20140702162934/http://www.jisc.ac.uk/whatwedo/programmes/digitisation.aspx>).
- ii (<https://www.hlf.org.uk/about-us/news-features/hlf-support-first-world-war-heritage>
- iii (<http://www.ahrc.ac.uk/documents/data/statistics/competition-statistics-annual-report-2013-14/>).
- iv (<https://www.hlf.org.uk/evaluation-guidance>)
- v (<http://microsites.oii.ox.ac.uk/tidsr/about-toolkit>).
- vi (<http://microsites.oii.ox.ac.uk/tidsr/case-studies>)
- vii (<http://microsites.oii.ox.ac.uk/tidsr/selected-references-toolkit>)
- viii (<http://microsites.oii.ox.ac.uk/tidsr/welcome>).
- ix (http://microsites.oii.ox.ac.uk/tidsr/sites/microsites.oii.ox.ac.uk.tidsr/files/7-43_SPHERE_Web%20survey%20report-FINAL.pdf).
- x Technically, bounce is a single 'engagement hit' of which page view is just one of six types.
- xi (<https://blog.kissmetrics.com/bounce-rate/?wide=1>).
- xii <https://www.kaushik.net/avinash/standard-metrics-revisited-3-bounce-rate/>

Appendix 1

Cymru1914 Online Survey

This survey is gathering feedback on the use of the Cymru1914 web site (<http://cymru1914.org/>) to help understand its use and users.

1. What is your age?
 - Under 18
 - 18-24
 - 25-34
 - 35-44
 - 45-54
 - 55-64
 - 65+
2. How often to you use the Cymru1914 web site?
 - Frequently (daily or weekly)
 - Occasionally (monthly, every 1-3 months)

-
- Infrequently (a few times a year or less)
 - Have not previously used the site
3. How did you become aware of the site?
- Mailing list
 - Personal email message
 - Reference from another source (e.g. web site, forum)
 - Search engine
 - Word of mouth
 - I was involved in the work of the site
 - Other (please specify):
4. Which user group do you belong to?
- Academic
 - Library & Archive Staff
 - Student (HE)
 - Schools (Teacher or Pupil)
 - Government
 - Museum Staff
 - General Public
 - Technical / e-Research
 - Creative Industries / Business
 - Media
 - Military
5. What purpose do you primarily use the Cymru1914 site for?
- Personal research
 - Schoolwork / homework
 - Teaching (primary / secondary)
 - Teaching (continuing / further / higher)
 - Academic research
 - Journalism / media / creative content
 - Community outreach / exhibitions
 - eGovernment
 - Reference (Library & Archive)
 - Other (please specify):
6. What subject area do you primarily use the Cymru1914 site for?
- Genealogy / Family History
 - Local History
 - Welsh History
 - Welsh Language / Culture
 - Military History / Commemoration
 - History of WWI
 - Tourism
 - Other (please specify):
7. As a result of using the site did you:
- Find the information you were looking for
 - Discover new sources
 - Change your view / opinion on a subject

-
- Become interested in a new subject
 - Recommend the site to someone else (If yes, see Q8)
 - Visit an online archive, library or museum
 - Visit a physical archive, library or museum
 - Visit a place mentioned on the site
8. If you recommended the site to anyone else, who?
- Family member
 - Friend
 - Local History / Genealogy Group member
 - Community group member
 - Teacher
 - Academic
 - Other (please specify):
9. Have you used the Social Media Links?
- No
 - Yes (select which): Twitter / Facebook / Google+
10. Have you downloaded any resources?
- No
 - Yes (select which): Newspapers / Archives & Manuscripts / Photographs / Journals / Sound
11. Overall, do you find use of the site:
Very Easy / Quite Easy / Quite Difficult / Very Difficult
Do you have any comments on your use?
12. Please rate the following aspects of the site on a scale of 1 – 5 or NA (not applicable)
(1 = Very unsatisfactory, 2 = Unsatisfactory, 3 = Satisfactory, 4 = Good, 5 = Excellent)
- Search Interface
 - Search Results
 - Image Zoom / Rotate
 - Further Information
 - Filter By
 - Browse Interface
 - Navigation
 - Page Layout
 - Site Information
13. Please indicate any other comments or questions you have about the site. If you would be willing to participate in a follow up interview or focus group, please leave your email address: